Mask-Compliant Orthogonal Precoding for Spectrally Efficient OFDM

Ravinder Kumar[1], Khawar Hussain[2] and Roberto López-Valcarce[2]

***Abstract*—Orthogonal precoding constitutes a powerful technique to reduce spectrum sidelobes of multicarrier signals. This reduction is bought at the cost of introducing precoder redundancy, which results in some throughput loss and additional precoding/decoding complexity. When the goal is to meet some spectral emission mask constraints, it is desirable to avoid unnecessary sidelobe suppression in order to keep precoder redundancy at a minimum. In this context, we introduce a general framework under which we develop a novel Lagrange multiplier-based mask-compliant orthogonal precoder design targeting minimal redundancy. We also adapt to this framework two previously proposed designs based on spectral notches and minimum out-of-band emission, respectively, to explicitly incorporate mask constraints. Simulation results are provided to show the effectiveness of the proposed designs under different practical masks for multicarrier wireless systems.**

I. Introduction

The quest for modulation techniques with ultra-flexible waveforms is still ongoing. The current and next generation wireless networks, including 4G Long Term Evolution (LTE), 5G New Radio (NR), and the cognitive radio (CR) paradigm, target a wide range of use cases as well as efficient utilization of spectral resources. To this extent, there is a need for a waveform that makes efficient use of the available spectrum for best quality of service while meeting diverse and strict spectrum emission requirements. These requirements are usually defined in terms of spectrum emission masks (SEM) or unwanted emission masks (UEM) [1]. Typically, the goal of such masks is to avoid interference to adjacent channels, although the SEM concept can be further extended to protect incumbent networks operating within the same band in typical spectrum sharing scenarios; such bands are termed primary user (PU) bands in CR networks [2], [3]. Most existing wireless technologies are subject to well-defined SEM requirements, e.g., CR networks [5], 5G NR [6], etc.

Orthogonal frequency division multiplexing (OFDM) has been regarded as the most favorable modulation scheme for (downlink) 4G, 5G and numerous other wireless technologies due to its robustness toward multipath fading channels, spectrum aggregation capabilities, good pairing with multiple-input multiple-output (MIMO) techniques, etc. [7], [8]. Despite numerous advantages, OFDM suffers from high sidelobe leakage due to the inherent sinc-like characteristics of each subcarrier waveform in the frequency domain. This poses a challenge for OFDM-based systems to meet SEM requirements while maximizing usage of available spectral resources. It becomes even more challenging in CR scenarios, where spectrum is shared with legacy PUs so that additional SEM constraints in PU bands are to be met. An exemplary SEM is shown in Fig.1, which illustrates the emission limits in both adjacent and PU bands. The power spectral density (PSD) of the signal generated by any wireless transmitter must be strictly below the corresponding SEM.

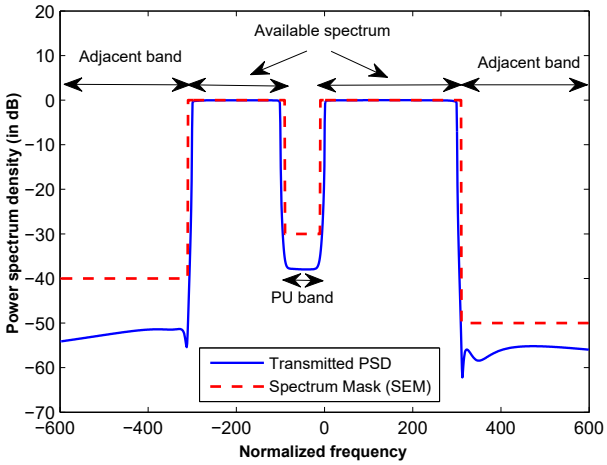

Fig. 1: Illustration of spectral emission mask (SEM) and mask-compliant power spectrum.

Although it is straightforward to reduce the impact of sidelobes by deactivating subcarriers (commonly known as *guard subcarriers*) at the band edges, this usually results in significant throughput loss, as the required number of guard subcarriers to meet SEM constraints can be large. Motivated by this, many sidelobe suppression techniques have been proposed, including time-domain methods such as filtering [9], windowing (or *pulse-shaping*) [10], [11], [12] and adaptive symbol transition (AST) [13]. Filtering usually requires filters with a long impulse response to avoid signal distortion; as this adds to the length of the channel impulse response, the effective guard interval of OFDM symbols is reduced. Windowing and AST have a similar effect, as part of the guard interval is used up for pulse-shaping purposes.

Another class of sidelobe suppression schemes operate in the frequency domain, such as active interference cancellation (AIC) [14], [15], [16], [17] and precoding techniques. Out of K available subcarriers, AIC methods explicitly reserve R

[1]R. Kumar is with Indian Institute of Technology, Roorkee, Uttarakhand, India.

[2]K. Hussain and R. López-Valcarce are with atlanTTic Research Center, University of Vigo, Spain. Their work is supported by Agencia Estatal de Investigación (Spain) and the European Regional Development Fund (ERDF) under projects WINTER (TEC2016-76409-C2-2-R, BES-2017-080305) and RODIN (PID2019-105717RB-C21), and by Xunta de Galicia (Agrupación Estratéxica Consolidada de Galicia accreditation 2020-2022).

of them (termed *cancellation carriers*) which are modulated as a function of the data symbols in the remaining data subcarriers in order to reduce unwanted emissions. At the receiver, cancellation carriers are simply discarded, and data subcarriers are demodulated as usual. In general, for a given SEM, AIC techniques are much more efficient than simply turning off guard subcarriers, in the sense that the number of cancellation carriers required for a given sidelobe suppression level is much smaller.

Spectral precoding can be regarded as a generalization of AIC in which there is no specific set of reserved subcarriers, with the number of data symbols per block $K-R$ smaller than or equal to the number of available subcarriers K to which they are mapped, so that $R \geq 0$ can be thought of as the *redundancy* of the precoder. For a given value of R, spectral precoding is generally more effective than AIC in terms of unwanted emission reduction, although the operation is not transparent to the receiver anymore, i.e., some in-band distortion is introduced, which may require appropriate decoding at the receiver side to avoid bit error rate (BER) degradation. Nevertheless, the potential of these techniques has attracted much interest, and many different precoder designs have been proposed [18], [19], [20], [21], [22], [23]. The notch-based precoder presented in [21] has redundancy $R=0$, and reduces sidelobes to a great extent by placing spectral nulls at some carefully selected out-of-band notch frequencies while minimizing in-band distortion as measured by the Error Vector Magnitude (EVM); the resulting distortion increases with the number of notch frequencies. The mask-compliant precoder (MCP) introduced in [24] generalizes the design from [21] with reduced EVM by avoiding unnecessary suppression at notch frequencies; however, [24] requires solving a convex optimization problem per block in an on-line fashion, which is computationally expensive especially in applications with a large number of subcarriers. A computationally more efficient alternative has been proposed in [25], but nevertheless the issue of on-line complexity remains. To address this issue, [26] considered mask-compliant *linear* precoders, in which the precoded symbols are obtained by multiplying original data symbols by a fixed matrix. This matrix was computed off-line to minimize EVM subject to the SEM constraints, so that on-line complexity is that of a $K \times K$ matrix-vector multiplication per data block, which can be further reduced if the structure of the precoding matrix is exploited [26].

All the aforementioned precoders have redundancy $R=0$, so that they trade off spectral performance for EVM degradation: the more stringent the SEM constraints, the worse the achieved EVM. Moreover, the achieved EVM is generally not uniform across subcarriers and tends to be worse at band edges, which translates in a degraded block error rate. A different tradeoff is possible in terms of throughput rather than EVM by using so-called *orthogonal* precoders, which have $R>0$ and use a fixed precoding matrix of size $K \times (K-R)$ with orthonormal columns [19], [22], [27]. Thanks to this orthogonality property, the receiver can easily invert the precoding operation without degrading the EVM; moreover, orthogonal precoding/decoding lend themselves to efficient on-line implementation (with complexity KR rather than K^2) via Householder block reflectors [28]. The orthogonal precoders from [19], [22] are based on a *total suppression criterion* (TSC) attempting to minimize the total out-of-band transmitted power, whereas the design from [27] is based on the notching criterion (NC), introducing spectral nulls at R carefully selected out-of-band frequencies analogously to [21]; the number and location of these notch frequencies directly determine the throughput and performance of the system. Extensions to a scenario with multiple cognitive users were proposed in [29].

None of these orthogonal precoding designs target SEM constraints specifically. Although it is true in general that they should be able to comply to a given mask by using a sufficiently large redundancy R, this may be inefficient, as in practice R should be kept as low as possible to avoid unnecessary throughput degradation. For that reason, in this paper we aim at designing orthogonal precoders capable of meeting SEM requirements with minimum redundancy. The proposed designs retain the desirable features of orthogonal techniques, i.e., no EVM degradation and powerful sidelobe suppression, while minimizing throughput loss and on-line implementation complexity. Specifically, our contributions can be summarized as:

- We present a unified framework/system model for mask compliant orthogonal precoding in terms of a unitary matrix and binary weight constraints.
- Based on this framework, we expose the structure of the optimal matrix and reformulate the design as a feasibility problem in terms of Lagrange multipliers.
- An algorithm for solving this problem is derived by exploiting the properties of the PSD as a function of the Lagrange multipliers.
- We apply the unified framework to develop NC- and TSC-based mask-compliant orthogonal designs targeting minimum redundancy.
- The performance of the proposed orthogonal designs is illustrated in a variety of scenarios including practical spectral masks from a number of wireless standards.

The organization of the paper is as follows: Sec. II presents a generalized system model for the study of orthogonal precoders with SEM constraints. A Lagrange multiplier-based mask-compliant orthogonal design is developed in Sec. III. The NC- and TSC-based orthogonal designs are revisited in Sec. IV, where modifications are proposed to accommodate SEM constraints. Computational complexity is discussed in Sec. V, and results are provided in Sec. VI. Finally, conclusions are drawn in Sec. VII.

Notation: vectors and matrices are respectively denoted by boldface lowercase and uppercase symbols. For a matrix $\mathbf{A}$, we denote its transpose, conjugate transpose, pseudoinverse, and trace by $\mathbf{A}^T$, $\mathbf{A}^H$, $\mathbf{A}^\dagger$ and $\mathrm{Tr}[\mathbf{A}]$, respectively. The 2-norm of vector $\mathbf{v}$ is denoted by $\|\mathbf{v}\|$. $\mathbf{I}_K$ denotes the $K \times K$ identity matrix, whereas $E[\cdot]$ is the expectation operator. The sinc function is defined as $\mathrm{sinc}(x)=\frac{\sin(\pi x)}{\pi x}$, and $\delta_{p,q}=1$ if $p=q$ and zero otherwise denotes Kronecker's delta.

II. General System Model

A unified system model for orthogonal precoding is presented in this section. Consider a multicarrier communication system using K subcarriers with indices in the set $\mathcal{K} = \{k_0, k_1, \ldots, k_{K-1}\}$. The $K \times 1$ vector of transmitted symbols is denoted by $\tilde{\mathbf{d}} = [\ \tilde{d}_{k_0} \quad \tilde{d}_{k_1} \quad \cdots \quad \tilde{d}_{k_{K-1}}\ ]^T$. Essentially, the k_i-th subcarrier is modulated by the complex symbol $\tilde{d}_{k_i}$. A general framework for the operation of orthogonal spectral precoders to obtain the precoded data vector $\tilde{\mathbf{d}}$ is stated as

$$\tilde{\mathbf{d}} = \mathbf{G}\mathbf{W}\mathbf{d} = \bar{\mathbf{G}}\bar{\mathbf{d}}, \tag{1}$$

where $\mathbf{G}$ is the $K \times K$ unitary precoding matrix, satisfying $\mathbf{G}^H\mathbf{G} = \mathbf{I}_K$, and $\mathbf{W}$ is a $K \times K$ diagonal weighting matrix with binary-valued diagonal elements $w_0, w_1, \ldots, w_{K-1} \in \{0, 1\}$. The precoding matrix is written columnwise as $\mathbf{G} = [\ \mathbf{g}_0 \quad \mathbf{g}_1 \quad \cdots \quad \mathbf{g}_{K-1}\ ]$, so that $\mathbf{g}_k$ denotes the k-th column of $\mathbf{G}$. In (1), $\mathbf{d} = [\ d_0 \quad d_1 \quad \cdots \quad d_{K-1}\ ]^T$ represents the $K \times 1$ vector of QAM data symbols. Whenever $w_k = 0$, the data symbol d_k on the k-th subcarrier is not transmitted[1], so that the number of nonzero weights effectively determines the system throughput. Let the number of zero weights (precoder redundancy) be denoted by R. The $K \times (K - R)$ matrix $\bar{\mathbf{G}}$ comprises those columns $\mathbf{g}_k$ corresponding to the indices for which $w_k = 1$, and the $(K - R) \times 1$ vector $\bar{\mathbf{d}}$ comprises the elements of $\mathbf{d}$ for the same indices.

Since $\bar{\mathbf{G}}$ has orthonormal columns, one has $\bar{\mathbf{G}}^H\bar{\mathbf{G}} = \mathbf{I}_{K-R}$, and the original symbols can be directly recovered from $\tilde{\mathbf{d}}$ as

$$\bar{\mathbf{d}} = \bar{\mathbf{G}}^H\tilde{\mathbf{d}}. \tag{2}$$

Moreover, this operation does not incur noise enhancement. Suppose a noisy version $\mathbf{r} = \tilde{\mathbf{d}} + \mathbf{n}$ is available at the receiver, with $\mathbf{n}$ a zero-mean noise vector with covariance $\mathbf{C} = E[\mathbf{n}\mathbf{n}^H]$. Then $\bar{\mathbf{G}}^H\mathbf{r} = \bar{\mathbf{d}} + \bar{\mathbf{G}}^H\mathbf{n}$, and the resulting noise component $\bar{\mathbf{G}}^H\mathbf{n}$ is zero-mean with covariance $\bar{\mathbf{G}}^H\mathbf{C}\bar{\mathbf{G}}$. By virtue of von Neumann's trace inequality [33] and the fact that $\bar{\mathbf{G}}^H\bar{\mathbf{G}} = \mathbf{I}_{K-R}$, one has

$$\begin{aligned} E\left[\|\bar{\mathbf{G}}^H\mathbf{n}\|^2\right] = \mathrm{Tr}[\bar{\mathbf{G}}^H\mathbf{C}\bar{\mathbf{G}}] &\leq \sum_{k=0}^{K-R-1} \lambda_k(\mathbf{C}) \\ &\leq \mathrm{Tr}[\mathbf{C}] = E\left[\|\mathbf{n}\|^2\right], \end{aligned} \tag{3}$$

where $\lambda_0(\mathbf{C}) \geq \lambda_1(\mathbf{C}) \geq \cdots \geq \lambda_{K-1}(\mathbf{C})$ are the eigenvalues of $\mathbf{C}$. Thus, (3) shows that the noise power is not increased by the decoding process. In particular, if the noise $\mathbf{n}$ is uncorrelated and with the same variance across the different subcarriers, i.e., $\mathbf{C} = \sigma^2\mathbf{I}_K$, then so is the noise component after decoding, since in that case $\bar{\mathbf{G}}^H\mathbf{C}\bar{\mathbf{G}} = \sigma^2\mathbf{I}_{K-R}$.

In order to derive the PSD of the transmitted signal, let the useful symbol and guard interval duration be T_s and T_g respectively (so that the subcarrier spacing is $\Delta f = \frac{1}{T_s}$). The cyclic-prefix based multicarrier signal corresponding to the transmission of a single data block can be written as

$$s(t) = \sum_{k \in \mathcal{K}} \tilde{d}_k e^{j2\pi \frac{k}{T_s} t} w(t), \tag{4}$$

where $w(t)$ is a rectangular windowing function given by $w(t) = 1$ for $-T_g \leq t < T_s$ and 0 otherwise. The Fourier transform of $s(t)$ is then given by

$$S(f) = \sum_{k \in \mathcal{K}} \tilde{d}_k a_k^*(f) = \mathbf{a}^H(f)\tilde{\mathbf{d}}, \tag{5}$$

where $\mathbf{a}(f) = [a_{k_0}(f), a_{k_1}(f), \ldots, a_{k_{K-1}}(f)]^T$, $a_k(f)$ is given by

$$a_k(f) = \mathrm{sinc}\left(T\left(f - \frac{k}{T_s}\right)\right) e^{j\pi(T_s - T_g)(f - \frac{k}{T_s})}, \tag{6}$$

and $T = T_s + T_g$ denotes the total symbol duration. Assuming independent and identically distributed symbols with zero mean and unit variance, so that $E[\mathbf{d}\mathbf{d}^H] = \mathbf{I}_K$, the PSD is given by [26]

$$\begin{aligned} P(f) &= E[|S(f)|^2] = \mathbf{a}^H(f)E[\mathbf{G}\mathbf{W}\mathbf{d}\mathbf{d}^H\mathbf{W}^H\mathbf{G}^H]\mathbf{a}(f) \\ &= \mathrm{Tr}[\mathbf{W}\mathbf{G}^H\mathbf{A}(f)\mathbf{G}], \qquad (7) \\ &= \sum_{k=0}^{K-1} w_k \mathbf{g}_k^H \mathbf{A}(f)\mathbf{g}_k, \qquad (8) \end{aligned}$$

where the $K \times K$ Hermitian matrix $\mathbf{A}(f) \triangleq \mathbf{a}(f)\mathbf{a}^H(f)$ has been introduced, and we have used the fact that $\mathbf{W}\mathbf{W}^H = \mathbf{W}$ due to the binary nature of the weights $\{w_k\}$.

Now let the SEM be given by $M(f)$, so that the transmitted PSD must satisfy

$$\mathrm{Tr}[\mathbf{W}\mathbf{G}^H\mathbf{A}(f)\mathbf{G}] \leq M(f) \quad \text{for all } f \in \mathbb{R}. \tag{9}$$

Note that the requirement (9) translates into infinitely many constraints, one for each value of $f \in \mathbb{R}$. In practice, it can be approximated by considering a finite number D of frequencies contained in the set $\mathcal{D} = \{f_0, f_1, \cdots, f_{D-1}\}$ [26] . In this way, the constraints (9) are replaced by

$$\sum_{k=0}^{K-1} w_k \mathbf{g}_k^H \mathbf{A}(f_m)\mathbf{g}_k \leq M(f_m), \quad \forall f_m \in \mathcal{D}. \tag{10}$$

In general, it suffices to construct the set $\mathcal{D}$ with frequencies in the out-of-band region $\mathcal{B}$ (which includes adjacent bands as well as coexisting PU bands) since orthogonal precoding schemes tend to preserve in-band flatness, which is usually the requirement of typical SEM specifications.

III. Lagrange Multiplier-based Orthogonal MCP

As existing orthogonal precoding techniques do not guarantee minimum redundancy R for a given set of SEM constraints, it is of interest to develop new orthogonal designs explicitly incorporating these requirements. The redundancy of the precoder can be written as $R = K - \sum_{k=0}^{K-1} w_k$, and the number of data symbols transmitted over the K subcarriers is $K - R$. Our goal is to design a mask-compliant precoder minimizing the redundancy R while retaining orthogonality, i.e., $\mathbf{G}^H\mathbf{G} = \mathbf{I}_K$. Note that minimizing R amounts to

[1]Note that such "not transmitted" data symbols, i.e., the entries of $\mathbf{d}$ not present in $\bar{\mathbf{d}}$, are just virtual symbols introduced here for notational convenience only. In practice the precoding operation is directly implemented as $\tilde{\mathbf{d}} = \bar{\mathbf{G}}\bar{\mathbf{d}}$.

maximizing the number (or equivalently, the sum) of nonzero weights. The corresponding optimization problem for finding $\mathbf{G}$ and $\mathbf{W}$ can be formulated as follows:

$$\mathbb{OP}1: \quad \max_{\mathbf{W},\mathbf{G}} \quad \sum_{k=0}^{K-1} w_k \tag{11}$$

subject to :

$$\sum_{k=0}^{K-1} w_k \mathbf{g}_k^H \mathbf{A}(f_m)\mathbf{g}_k \leq M(f_m), \quad m = 0, 1, \ldots, D-1, \tag{12}$$

$$\mathbf{g}_p^H \mathbf{g}_q = \delta_{p,q}, \quad p, q = 0, 1, \ldots K-1, \tag{13}$$

$$w_k(1 - w_k) = 0, \quad k = 0, 1, \ldots, K-1. \tag{14}$$

The objective function in (11) can be regarded as a measure of throughput, which is to be maximized. The SEM requirement is given by (12); the constraint (13) ensures orthogonality, whereas the binary weight constraint is enforced by (14).

$\mathbb{OP}1$ belongs to the class of mixed-integer nonlinear problems (MINLPs), and is non-convex due to constraint (12). A number of well-established and effective techniques are available for *convex* MINLPs, often based on continuous relaxation of the integer-valued constraints, which yields a convex optimization problem, see e.g. [30]. Non-convex MINLPs, in contrast, are regarded as much more difficult to deal with, as even after continuous relaxation the problem remains non-convex. In our case, the cubic non-convex constraints (12) can be reformulated as quadratic constraints by introducing additional variables, so that in principle the numerical techniques of [31, Sec. 5] should be applicable in order to find an approximate solution. We shall not pursue such approach here, however, as it is likely to lead to significant complexity without providing insights on the nature of the solution, and leave its study for future work. Instead, we find first the structure of the optimal precoding matrix for fixed $\mathbf{W}$; then, an iterative method is proposed for finding suitable weights so as to approximately maximize throughput.

A. Structure of optimal $\mathbf{G}$

To expose the structure of the optimal precoding matrix, consider the Lagrangian of $\mathbb{OP}1$, which can be written as

$$\begin{aligned}\mathcal{L}(\mathbf{W}, \mathbf{G}, \boldsymbol{\nu}, \boldsymbol{\mu}, \boldsymbol{\gamma}) &= \sum_{k=0}^{K-1} w_k + \sum_{p=0}^{K-1}\sum_{q=0}^{K-1} \nu_{p,q}(\mathbf{g}_p^H \mathbf{g}_q - \delta_{p,q}) \\ &\quad + \sum_{m=0}^{D-1} \mu_m \left(M(f_m) - \sum_{k=0}^{K-1} w_k \mathbf{g}_k^H \mathbf{A}(f_m)\mathbf{g}_k \right) \\ &\quad + \sum_{k=0}^{K-1} \gamma_k w_k(1 - w_k),\end{aligned} \tag{15}$$

where μ_m, γ_k and $\nu_{p,q}$ are Lagrange multipliers, and $\mu_m \geq 0$ for $m = 0, 1, \ldots, D-1$.

The optimal $\mathbf{G}$ and $\mathbf{W}$ should maximize the Lagrangian (15) and satisfy the respective orthogonality and binary constraints, when the Lagrange multipliers take on their optimal values. Note that, for a given set of multipliers $\{\mu_m\}$, and upon defining the positive semidefinite matrix

$$\mathbb{A} \triangleq \sum_{m=0}^{D-1} \mu_m \mathbf{A}(f_m), \tag{16}$$

it is seen that the Lagrangian (15) is maximized w.r.t. $\mathbf{G}$ unitary if $f(\mathbf{G}) \triangleq \sum_{k=0}^{K-1} w_k \mathbf{g}_k^H \mathbb{A} \mathbf{g}_k$ is minimized. Since the weights w_k are binary valued, it follows that $f(\mathbf{G}) = \mathrm{Tr}[\bar{\mathbf{G}}^H \mathbb{A} \bar{\mathbf{G}}]$, where $\bar{\mathbf{G}}$ comprises the $K - R$ columns of $\mathbf{G}$ corresponding to the indices for which $w_k = 1$. Therefore, $f(\mathbf{G})$ is minimized when the columns of $\bar{\mathbf{G}}$ constitute an orthonormal basis for the subspace spanned by the $K - R$ eigenvectors corresponding to the $K - R$ smallest eigenvalues of $\mathbb{A}$ [32]. The R remaining columns of $\mathbf{G}$ must therefore span the eigensubspace corresponding to the R dominant eigenvalues. Hence, without loss of generality, the columns of the optimal $\mathbf{G}$ can be taken as the eigenvectors of $\mathbb{A}$. The eigenvalues are denoted as $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_{K-1}$ and arranged in the diagonal matrix $\boldsymbol{\Lambda}$, so that the eigenvalue decomposition (EVD) of $\mathbb{A}$ can be written as $\mathbb{A} = \mathbf{G}\boldsymbol{\Lambda}\mathbf{G}^H$.

B. Choosing weights w_k

With the opimal unitary precoding matrix $\mathbf{G}$, the Lagrangian (15) reduces to

$$\begin{aligned}\mathcal{L} &= \sum_{k=0}^{K-1} w_k(1 - \lambda_k) \\ &\quad + \sum_{m=0}^{D-1} \mu_m M(f_m) + \sum_{k=0}^{K-1} \gamma_k w_k(1 - w_k).\end{aligned} \tag{17}$$

The optimal choice of weights $w_k \in \{0, 1\}$ should maximize (17). The last two terms in (17) are constant for binary valued weights, and since $1 - \lambda_0 \leq 1 - \lambda_1 \leq \cdots \leq 1 - \lambda_{K-1}$, it is clear that the optimal set of weights is

$$w_k = \begin{cases} 0, & 0 \leq k \leq R-1, \\ 1, & R \leq k \leq K-1, \end{cases} \tag{18}$$

which is consistent with the fact that the columns of $\mathbf{G}$ corresponding to $w_k = 1$ must span the subspace associated to the $K - R$ smallest eigenvalues, as found above.

Now the problem reduces to finding the smallest R for which a combination of non-negative multipliers μ_m satisfying the mask constraint (12) exists. Hence, for fixed R, we can pose a feasibility problem as

$$\mathbb{OP}2: \quad \text{Find } \{\mu_m\} \tag{19}$$

subject to :

$$\sum_{k=R}^{K-1} \mathbf{g}_k^H \mathbf{A}(f_m)\mathbf{g}_k \leq M(f_m), \quad m = 0, 1, \ldots, D-1, \tag{20}$$

$$\mu_m \geq 0, \quad m = 0, 1, \ldots, D-1. \tag{21}$$

Note that the $\mathbf{g}_k$'s in (20) are implicit functions of $\{\mu_m\}$, since they are eigenvectors of the matrix $\mathbb{A}$ given in (16). The goal is to find the smallest R for which $\mathbb{OP}2$ is feasible.

C. Solving the feasibility problem $\mathbb{OP}2$ for fixed R

In order to understand the impact of each individual multiplier μ_m on the spectrum at different frequencies, the following theorem will be useful. Its proof can be found in the Appendix.

Theorem 1: If the eigenvalues of $\mathbb{A}$ are all distinct, then:

1) The PSD at frequency f_m, i.e., $P(f_m) = \sum_{k=R}^{K-1} \mathbf{g}_k^H \mathbf{A}(f_m)\mathbf{g}_k$, is a non-increasing function of μ_m.
2) The weighted sum of PSD values at all frequencies in $\mathcal{D}$ except f_m, defined as $Q(f_{\not m}) \triangleq \sum_{n\neq m} \mu_n P(f_n)$, is a non-decreasing function of μ_m.
3) The weighted sum of PSD derivatives with respect to a given μ_m is zero, i.e., $\sum_{n=0}^{D-1} \mu_n \frac{\partial P(f_n)}{\partial \mu_m} = 0$.

These properties allow an interpretation of multiplier changes as "force fields" acting at different frequencies of the power spectrum: increasing the multiplier μ_m will try to force the PSD downwards at frequency f_m, while pulling up the weighted spectrum at the remaining frequencies in $\mathcal{D}$. Based on these facts, a heuristic iterative algorithm for solving $\mathbb{OP}2$ is presented as Algorithm 1.

Algorithm 1 Solving feasibility problem $\mathbb{OP}2$ for fixed R

Set R, $\mathcal{D} = \{f_0, \ldots, f_{D-1}\}$, $\beta > 0$, S, p
Initialize $i = 1$, `Stop` = False, $\mu_m^{(0)}$ $\forall m \in \{0, 1, \ldots, D-1\}$
while `Stop` = False **do**
  Evaluate $\mathbb{A}^{(i)} = \sum_{m=0}^{D-1} \mu_m^{(i-1)} \mathbf{A}(f_m)$
  Find the unitary precoding matrix $\mathbf{G}^{(i)}$ through the EVD $\mathbb{A}^{(i)} = \mathbf{G}^{(i)}\mathbf{\Lambda}^{(i)}(\mathbf{G}^{(i)})^H$
  Evaluate $P^{(i)}(f_m) = \sum_{k=R}^{K-1} (\mathbf{g}_k^{(i)})^H \mathbf{A}(f_m)\mathbf{g}_k^{(i)}$ $\forall f_m \in \mathcal{D}$
  if $P^{(i)}(f_m) \leq M(f_m), \forall f_m \in \mathcal{D}$ **then**
    `Stop` $\leftarrow$ True
    return *feasible*, $\mathbf{G}^{(i)}$
  end if
  for $m = 0, 1, \ldots, D-1$ **do**
    Initialize $J^{(i)} = 0$
    if $P^{(i)}(f_m) > M(f_m)$ **then**
      Update step: $\widetilde{\mu}_m^{(i)} = (1+\beta)\mu_m^{(i-1)}$
      $J^{(i)} \leftarrow J^{(i)} + P^{(i)}(f_m) - M(f_m)$
    end if
  end for
  Find $q = \arg\max_m \left\{\widetilde{\mu}_m^{(i)}\right\}$
  for $m = 0, 1, \ldots, D-1$ **do**
    Normalization step: $\mu_m^{(i)} = \widetilde{\mu}_m^{(i)} / \widetilde{\mu}_q^{(i)}$
  end for
  if $J^{(i-j)} > J^{(i-j-1)}$ at least p times for $j \in \{0, 1, \ldots, S-1\}$ **then**
    `Stop` $\leftarrow$ True
    return *infeasible*
  end if
  $i \leftarrow i + 1$
end while

At each iteration, the frequency points of $\mathcal{D}$ at which the PSD is above the SEM are identified and, in order to push the PSD downwards at such points, the corresponding multipliers are increased by a factor $1 + \beta$, where $\beta > 0$ is small. After this step, all multipliers are normalized so that the maximum multiplier after normalization is equal to 1. The effect of this normalization step is twofold: first, it avoids having runaway multipliers as iterations progress, and second, it tends to decrease those multipliers corresponding to frequencies in $\mathcal{D}$ at which the SEM constraint is loosely satisfied, potentially providing additional margin to increase the multipliers corresponding to frequencies at which the SEM constraints are violated.

The parameter β controls the convergence rate of the algorithm: it must be sufficiently small to allow for smooth progress, yet not too small in order to avoid very slow convergence. If at any iteration all SEM constraints are satisfied, the algorithm terminates and returns the feasible precoder.

On the other hand, the problem is declared infeasible as soon as it is detected that the sum of all positive differences between the PSD and the SEM has increased at least p times during the last S iterations. This heuristic rule quickly spots infeasible cases, in which typically after a transient decay stage such sum eventually starts to increase or oscillate. By judicious choice of S and p, borderline feasible cases in which such sum does not necessarily decay monotonically are not prematurely stopped. In practice, $S = 5$ and $p = 3$ turn out to provide a good tradeoff between these two effects.

We shall refer to this method as *Lagrange multiplier (LM)-based design.* Proper initialization of the multipliers in Algorithm 1 is important from a computational complexity point of view and is discussed next.

D. Finding the minimum redundancy

The final step is to find the minimum value of R for which problem $\mathbb{OP}2$ is feasible. A straightforward approach is to test feasibility of $\mathbb{OP}2$ for all possible values of R sequentially, starting with $R = 0$, then $R = 1$, etc., and stopping as soon as a feasible solution is found. However, this can be computationally expensive, since declaring a given value of R as infeasible often implies performing a larger number of iterations in Algorithm 1. Therefore, it is preferable, in terms of computational cost, to operate in the opposite direction: a sufficiently large value of R is first selected for which there is high confidence that a feasible solution can be found, and then the value of R is sequentially decreased. The advantage is that the number of iterations of Algorithm 1 that are needed to find a feasible solution when one exists will be much smaller than those required to declare infeasibility (especially when R is large). Additionally, when the (large) initial value of R is considered, we propose to initialize the Lagrange multipliers in Algorithm 1 as being all equal[2]. As the redundancy is successively reduced, the initial values of the multipliers for the current value of R are taken as the final values obtained in the previous step, i.e., for redundancy $R+1$. This *multiplier propagation* strategy results in fewer iterations than initializing all multipliers afresh for each R.

[2]Note that any initial value $\mu_m^{(0)} = \mu > 0$ can be chosen, as it only affects the eigenvalues of $\mathbb{A}^{(1)}$, but not its eigenvectors. For convenience, we have chosen $\mu = 1$.

IV. MCP BASED ON EXISTING ORTHOGONAL PRECODING TECHNIQUES

In this section we briefly describe two previous orthogonal precoding techniques, and place them under the generalized framework of Sec. II. Appropriate modifications are then presented in order to generate SEM-compliant orthogonal precoders based on these designs.

A. Notching criterion (NC) based precoder

The NC-based orthogonal precoder design [27] selects a number of *notch frequencies* at which the PSD is forced to be zero. Let the set of notch frequencies be $\mathcal{R} = \{f_0,\ f_1,\ \cdots, f_{R-1}\}$, where $\mathcal{R} \subseteq \mathcal{D}$, so that the goal is to have $P(f_m) = 0\ \forall f_m \in \mathcal{R}$. These notching conditions imply that the precoder must lie in a $(K-R)$-dimensional subspace, i.e., R denotes again the redundancy in the precoding operation. We can write

$$P(f) = \sum_{k=R}^{K-1} \mathbf{g}_k^H \mathbf{A}(f) \mathbf{g}_k = \mathrm{Tr}[\mathbf{W}\mathbf{G}^H \mathbf{A}(f) \mathbf{G}], \tag{22}$$

with weights $w_k = 0$ for $k = 0, 1, \ldots, R-1$ and $w_k = 1$ for $k = R, R+1, \ldots, K-1$. The notching conditions imply that

$$\sum_{f_m \in \mathcal{R}} P(f_m) = \mathrm{Tr}[\mathbf{W}\mathbf{G}^H \mathbb{A}_{\mathrm{NC}} \mathbf{G}] = 0, \tag{23}$$

where $\mathbb{A}_{\mathrm{NC}} \triangleq \sum_{f_m \in \mathcal{R}} \mathbf{A}(f_m)$ is a rank-R matrix. Note that (23) is satisfied by taking $\mathbf{G}$ from the EVD of $\mathbb{A}_{\mathrm{NC}}$, with the columns $\mathbf{g}_0, \ldots, \mathbf{g}_{R-1}$ corresponding to the R nonzero eigenvalues.

The choice of number and positions of notch frequencies to achieve desired spectral characteristics in NC-based precoding remains an open problem. When the goal is to satisfy the SEM constraints, we propose a heuristic approach for the selection of suitable notch frequencies, summarized in Algorithm 2. This approach starts with an empty set of notch frequencies; a new element at a time is iteratively added to this set by selecting the frequency from a prespecified set $\mathcal{D}$ at which the PSD is *maximally* above the SEM. The process continues until all SEM constraints are satisfied.

Algorithm 2 NC-based MCP

1: Set $\mathcal{D} = \{f_0, \ldots, f_{D-1}\}$
2: Initialize $\mathcal{R} = \emptyset$, $\mathbf{G} = \mathbf{I}_K$, $\mathbb{A}_{\mathrm{NC}} = \mathbf{0}$
3: **for** $R = 0, 1, 2, \ldots$ **do**
4: Evaluate $P(f_m) = \sum_{k=R}^{K-1} \mathbf{g}_k^H \mathbf{A}(f_m) \mathbf{g}_k\ \forall f_m \in \mathcal{D}$
5: **if** $P(f_m) \leq M(f_m), \forall f_m \in \mathcal{D}$ **then**
6: return *feasible*, R, $\mathbf{G}$
7: **else**
8: Find $n = \arg\max_m \{P(f_m) - M(f_m)\}$
9: Update $\mathcal{R} \leftarrow \mathcal{R} \cup \{f_n\}$ and $\mathbb{A}_{\mathrm{NC}} \leftarrow \mathbb{A}_{\mathrm{NC}} + \mathbf{A}(f_n)$
10: Find precoding matrix through EVD as $\mathbb{A}_{\mathrm{NC}} = \mathbf{G}\mathbf{\Lambda}\mathbf{G}^H$
11: **end if**
12: **end for**

B. Total Suppression Criterion (TSC) based Precoder

Rather than placing nulls at the prespecified frequencies as in the NC-based design, TSC-based precoders [19], [22] minimize the total out-of-band radiation leakage, i.e., the goal is to minimize

$$\int_{f \in \mathcal{B}} P(f)\ df = \mathrm{Tr}[\mathbf{W}\mathbf{G}^H \mathbb{A}_{\mathrm{TSC}} \mathbf{G}], \tag{24}$$

where $\mathcal{B}$ is the out-of-band frequency region, and $\mathbb{A}_{\mathrm{TSC}} \triangleq \int_{f \in \mathcal{B}} \mathbf{A}(f) df$. Again, the diagonal matrix $\mathbf{W}$ contains R diagonal entries equal to zero, and the remaining $K-R$ entries are equal to 1. It follows that (24) is minimized when $\mathbf{G}$ comprises the eigenvectors of $\mathbb{A}_{\mathrm{TSC}}$ and the R null weights in $\mathbf{W}$ correspond to the largest R eigenvalues [32]. As in any orthogonal precoder design, the number of null weights, i.e., the redundancy R, determines a tradeoff between throughput degradation and performance, which in this case is measured in terms of total out-of-band radiation.

Note that the TSC-based precoder $\mathbf{G}$ does not depend on the selected value of R; rather, this value just determines how many columns of $\mathbf{G}$ are selected. Therefore, when the final goal is to comply to a spectral mask, a direct search method to minimize redundancy can be adopted by just sequentially checking if the SEM constraints are met for $R = 0, 1, 2, \ldots$ until a feasible solution is found.

V. COMPLEXITY ANALYSIS

Computation of the precoding matrix $\mathbf{G}$ can be done offline, and it has to be performed just once as long as the SEM and subcarrier allocation does not change. On the other hand, there are systems that use varying portions of the frequency band, so that $\mathbf{G}$ needs to be recomputed from time to time. The three methods discussed in Secs. III and IV require the computation of the EVD of certain $K \times K$ matrices, and the computational cost of each EVD is $O(K^3)$. Thus, offline complexity can be measured by the number of such EVDs required for each method. For example, the TSC-based design of Sec. IV-B is the computationally cheapest of the three methods discussed in this paper, as it requires a single EVD. For the NC-based design of Sec. IV-A, the number of EVDs equals the redundancy R of the obtained solution, which is problem-dependent; thus, in settings with tighter SEMs, in which the necessary redundancy can be expected to be larger, the offline complexity will correspondingly increase.

For the LM-based mask-compliant design of Sec. III, the number of EVDs is roughly equal to the total number of iterations required until a nonfeasible value of R is found, which depends on the particular scenario. Again, settings with more stringent SEM requirements are likely to take more iterations, but it does not seem possible to determine or even upper bound this iteration count in general. Choosing a larger parameter β in Algorithm 1 will generally speed up convergence, but too large a value may destabilize the algorithm; we have found that $\beta = 0.4$ provides a good tradeoff between convergence speed and stability.

Once the precoding matrix is obtained, the online computational load is due to the computation of (1) at the transmitter,

and (2) at the receiver, and is the same for any orthogonal precoder design. Implementing these by direct matrix-vector multiplication operation requires $\mathcal{O}(K(K-R)) \approx \mathcal{O}(K^2)$ complexity. This can be substantially reduced by adopting the Householder reflector-based implementation proposed in [28] resulting in $\mathcal{O}(KR)$ complexity, which further illustrates the benefits of having as small a redundancy R as possible.

VI. Simulation Results

In this section, we evaluate the performance of the proposed techniques by considering several test and practical SEMs. The following scenarios are studied.

A. Scenario 1: Symmetric mask

In the first setup we assume a multicarrier system with RF bandwidth $W = 10$ MHz, K active subcarriers, and subcarrier spacing $\Delta f = \frac{1}{T_s} = \frac{W}{K+1}$ MHz. The set of subcarrier indices is $\mathcal{K} = \{-\frac{K}{2}, \ldots, -1, 1, \ldots, \frac{K}{2}\}$ (the DC subcarrier is explicitly turned off). A piecewise flat and symmetric test mask is considered:

$$\text{Mask 1:} \quad M(f) = \begin{cases} 1, & |f| \leq 5 \text{ MHz}, \\ M_1, & 5 \text{ MHz} < |f| \leq 10 \text{ MHz}. \end{cases} \tag{25}$$

A discrete frequency grid $\mathcal{D}$ is obtained by uniformly sampling $M(f)$ every $\frac{\Delta f}{2}$ Hz in the out-of-band region 5 MHz $< |f| \leq$ 10 MHz, i.e., $\mathcal{D} = \left\{ \pm \left(\frac{W}{2} + n\frac{\Delta f}{2} \right) \mid n = 1, 2, \cdots, K \right\}$. Mask 1 corresponds to a typical symmetric scenario where out-of-band radiation in adjacent channels is to be limited to $10 \log_{10} M_1$ dB below the in-band transmission level.

Table I shows the precoding redundancy $R = K - \sum_i w_i$ obtained with the orthogonal precoding techniques of Secs. III, IV-A and IV-B for different values of M_1, assuming $K = 512$ and a cyclic prefix length $T_g = \frac{T_s}{16}$. As expected, as mask requirements become tighter, precoding redundancy must increase, which in turn degrades system throughput. Comparing the three different techniques, the LM-based MCP design achieves smaller redundancy, and therefore better spectrum utilization, than the NC-based and TSC-based MCP approaches.

The corresponding PSDs for $M_1 = -60$ dB are shown in Fig. 2, together with that of a fully-loaded unprecoded system, for reference. It is seen that the LM-based technique avoids unnecessary suppresion of out-of-band radiation, providing just enough attenuation to comply to the SEM requirements. Table II shows the number of iterations run by Algorithm 1 for $M_1 = -60$ dB and different redundancy values, using a factor $\beta = 0.4$ and stopping parameters $S = 5$, $p = 3$. For $R \geq 16$, a mask-compliant precoder is obtained in the first iteration, whereas for $R = 12$ it took 4 iterations to declare the problem infeasible, terminating the procedure. Assuming an initial value $R = 20$, the whole process took a total of 22 EVDs[3].

Table III shows the spectrum utilization achieved with the three mask-compliant orthogonal designs for $M_1 = -50$ dB, and for various CP overheads T_g/T_s and number active subcarriers K (the bandwidth W is 10 MHz throughout, so that varying K amounts to varying the subcarrier spacing $\Delta f = \frac{W}{K+1}$). It is seen that for all designs and for a given useful symbol duration $T_s = \frac{1}{\Delta f}$, spectral utilization somewhat degrades as the CP overhead increases, although this effect is not very significant unless the number of active subcarriers is small. For a given CP overhead, spectral utilization improves with larger K (smaller subcarrier spacing).

[3]Taking into account that, with multiplier propagation, the EVD for the first iteration with a given value of R can be directly obtained from that of the last iteration with the previous value $R + 1$.

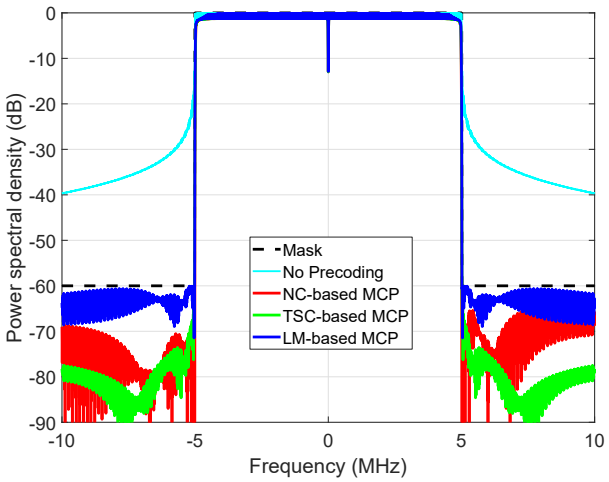

Fig. 2: PSD of various designs for Mask 1 with $M_1 = -60$ dB. $K = 512$, CP overhead $1/16$.

M_1 (dB)	-30	-40	-50	-60	-70	-80
NC-based	6	10	12	16	18	21
TSC-based	8	11	15	17	20	22
LM-based	6	8	11	13	16	18

TABLE I: Redundancy R of various mask-compliant orthogonal precoding designs for Mask 1. $K = 512$, CP overhead $1/16$.

R	20	$\cdots$	16	15	14	13	12
Iterations	1	$\cdots$	1	4	6	11	4
Feasible?	✓	$\cdots$	✓	✓	✓	✓	✗

TABLE II: No. of iterations in Algorithm 1 ($\beta = 0.4$, $S = 5$, $p = 3$) for Mask 1 with $M_1 = -60$ dB. $K = 512$, CP overhead $1/16$.

B. Scenario 2: Non-symmetric mask

In the second scenario we consider a similar setting to that from Scenario 1, but with a mask with different requirements in the left and right adjacent channels:

$$\text{Mask 2:} \quad M(f) = \begin{cases} M_{\text{left}}, & -10 \text{ MHz} \leq f < -5 \text{ MHz}, \\ 1, & |f| \leq 5 \text{ MHz}, \\ M_{\text{right}}, & 5 \text{ MHz} < f \leq 10 \text{ MHz}. \end{cases} \tag{26}$$

Table IV compares precoding redundancy for different precoding techniques for Mask 2, assuming $M_{\text{left}} = -60$ dB and with different values of M_{right}, for a system with $K = 512$

K	$\frac{T_g}{T_s}$ \	1/32			1/16			1/8			1/4		
		LM	NC	TSC	LM	NC	TSC	LM	NC	TSC	LM	NC	TSC
32		78.13	75	75	75	75	71.87	75	68.75	68.75	75	68.75	68.75
64		87.50	85.94	84.38	85.94	84.38	84.38	85.94	84.38	82.81	85.94	84.38	84.38
128		92.97	92.19	91.41	92.19	92.19	90.62	92.19	91.41	89.84	92.19	90.63	90.62
256		96.09	95.31	94.92	96.09	95.7	94.53	96.09	95.31	94.53	96.09	95.31	94.53
512		98.05	97.46	97.07	97.85	97.66	97.07	97.85	97.27	97.07	97.85	97.66	97.26
1024		98.93	98.63	98.53	98.93	98.63	98.44	98.83	98.63	98.34	98.83	98.63	98.44

TABLE III: Spectral utilization $\frac{K-R}{K} \times 100$ under Mask 1 for $M_1 = -50$ dB.

M_{right} (dB)	-30	-40	-50	-60	-70	-80
NC-based	11	13	14	16	17	19
TSC-based	17	17	17	17	20	22
Proposed	9	10	11	12	14	15

TABLE IV: Redundancy R of various mask-compliant orthogonal precoding designs for Mask 2, with $M_{\text{left}} = -60$ dB. $K = 512$, CP overhead $1/16$.

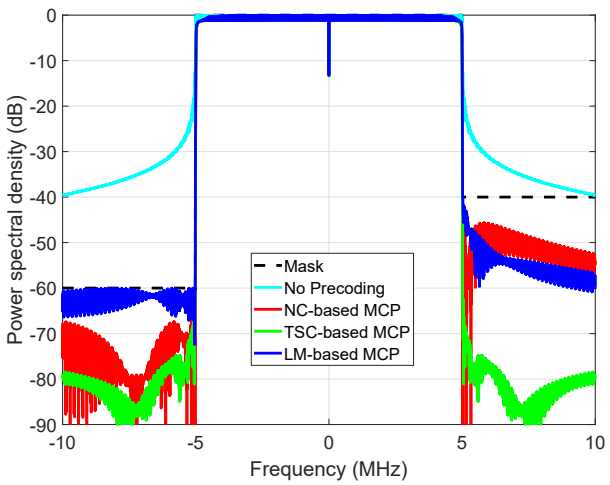

Fig. 3: PSD of various designs for Mask 2 with $M_{\text{left}} = -60$ dB, $M_{\text{right}} = -40$ dB. $K = 512$, CP overhead $1/16$.

subcarriers and $1/16$ CP overhead. The LM-based design from Sec. III (using $\beta = 0.4$, $S = 5$, $p = 3$) consistently achieves the smallest precoding redundancy, followed by the NC-based and TSC-based designs from Secs. IV-A and IV-B, respectively. Note in particular that the TSC-based design is oblivious to the fact that the mask is non-symmetric, and thus it is outperformed by the other two designs as it cannot exploit the fact that SEM requirements are less strict in one of the adjacent bands. The NC-based design is sufficiently flexible to exploit lack of symmetry in the SEM, but nevertheless it is outperformed by the LM-based design. The number of iterations for the latter in this scenario is reported in Table V; with an initial value $R = 20$, the whole process took 41 EVDs altogether. Fig. 3 shows the corresponding PSDs, for $M_{\text{left}} = -60$ dB and $M_{\text{right}} = -40$ dB.

R	20	$\cdots$	16	15	14	13	12	11	10
Iterations	1	$\cdots$	1	4	5	5	9	14	9
Feasible?	✓	$\cdots$	✓	✓	✓	✓	✓	✓	✗

TABLE V: No. of iterations in Algorithm 1 ($\beta = 0.4$, $S = 5$, $p = 3$) for Mask 2 with $M_{\text{left}} = -60$ dB, $M_{\text{right}} = -40$ dB. $K = 512$, CP overhead $1/16$.

C. Scenario 3: In-band SEM requirements

The mask for this scenario incorporates requirements for coexistence with PU subbands, along with non-symmetric requirements for adjacent bands. The setting is similar to that in Scenario 2, but with the transmission channel being shared with three primary users with different mask constraints. From the original set of subcarrier indices $\{-256, \ldots, -1, 1, \ldots, 256\}$, the following subcarrier ranges are turned off: $\{-160, \ldots, -129\}$ for PU 1 (32 subcarriers), $\{-64, \ldots, -33\}$ for PU 2 (32 subcarriers), and $\{65, \ldots, 128\}$ for PU 3 (64 subcarriers). The corresponding mask levels are -30, -40 and -50 dB for PU 1, 2 and 3, respectively. In this way, there remain $K = 384$ subcarriers available for data transmission.

Fig. 4 shows the PSD obtained with the three considered designs. The redundancy needed to satisfy the SEM requirements is $R = 41$, $R = 35$ and $R = 28$ for the TSC-based, NC-based and LM-based designs, corresponding to spectral utilization of 89.32%, 90.88% and 92.7%, respectively. It is seen that the TSC-based design has a hard time meeting the mask constraints at the band edges, which results in unnecessarily low emission levels farther away from such edges. The other two schemes, and in particular the LM-based design, are able to better adjust to the SEM characteristics.

D. Other SEMs

In this section we consider several practical cases corresponding to a number of wireless standards. First, consider the IEEE 802.22 standard for wireless regional area network (WRAN) using white spaces in the television (TV) frequency spectrum [5]. Assuming 6-MHz TV channels, the subcarrier spacing is $\Delta f = 3.348$ kHz, and the standard contemplates the usage of subcarriers with indices $\mathcal{K} = \{-840, \ldots, -1, +1, \ldots, +840\}$, resulting in an occupied bandwidth of 5.625 MHz. The out-of-band emission requirements issued by the Federal Communications Commission

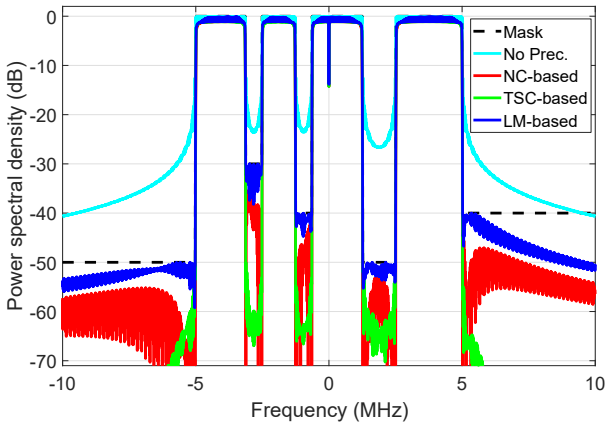

Fig. 4: PSD of various designs with in-band PU subbands to protect (Scenario 3).

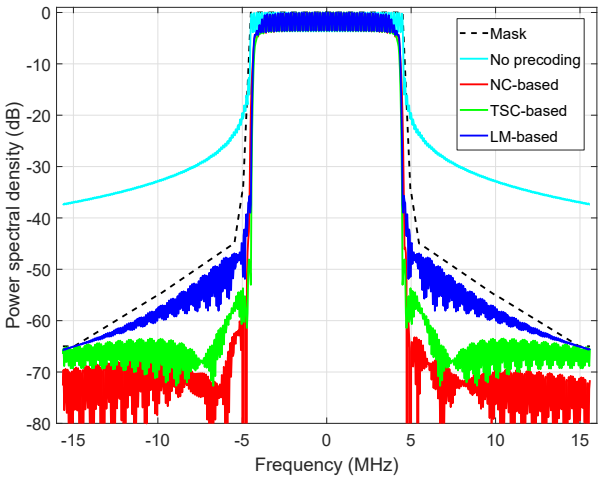

Fig. 6: PSD of various designs. IEEE 802.11p Class D mask.

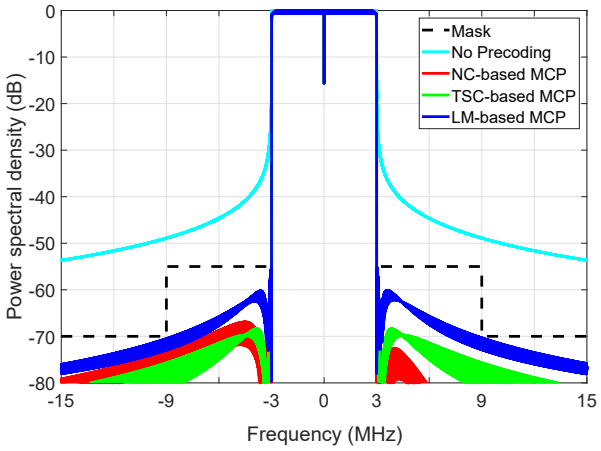

Fig. 5: PSD of various designs. IEEE 802.22 mask.

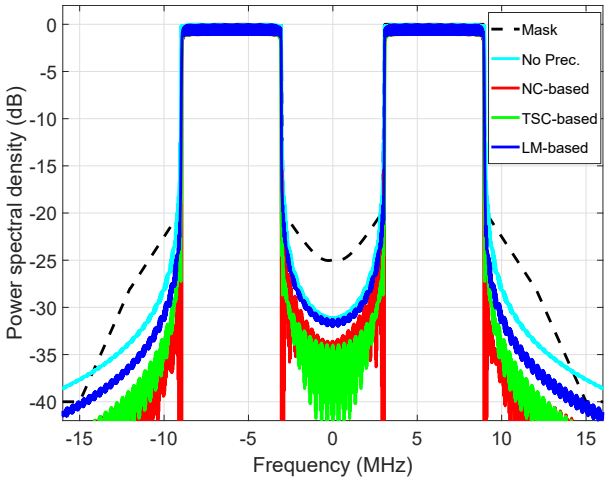

Fig. 7: PSD of various designs. IEEE 802.11af mask.

(FCC) for TV band devices translate into a 55.4 dB drop of the PSD in the first adjacent channels, and 68.52 dB in the second adjacent channels (assuming fixed devices with 36 dBm transmitted in-channel EIRP) [34]. In this framework, and assuming $1/32$ CP overhead, the orthogonal designs are able to meet these requirements with redundancy $R = 8$ (TSC-based), $R = 10$ (NC-based) and $R = 6$ (LM-based). However, the use of precoding allows to increase the number of used subcarriers over the channel bandwidth. Consider for instance the extreme case in which $\mathcal{K} = \{-895, \ldots, -1, +1, \ldots, +895\}$ resulting in an occupied bandwidth of 5.996 MHz. Then the out-of-band emission requirements can be met with redundancy $R = 19$ (TSC-based), $R = 18$ (NC-based) and $R = 14$ (LM-based), as shown in Fig. 5. These larger redundancy values still pay off because of the larger number of available subcarriers.

As a second example, consider the IEEE 802.11p Wireless LAN standard for Wireless Access in Vehicular Environments [35] with 10-MHz channelization. The standard contemplates the use of 52 subcarriers with 156.25-kHz subcarrier spacing and $1/4$ CP overhead, for an occupied bandwidth of 8.125 MHz. Four different SEMs are specified depending of the maximum output power, with the most stringent one corresponding to Class D (28.8 dBm) devices. The requirements of the Class D SEM in this setting can be met using the orthogonal designs, with redundancy $R = 7$ (TSC-based), $R = 8$ (NC-based) and $R = 6$ (LM-based) . Again, a more aggressive use of spectrum is possible: for example, if 58 subcarriers are used, the orthogonal designs are still able to meet the SEM requirements, this time with redundancy $R = 10$ (TSC-based), $R = 10$ (NC-based) and $R = 8$ (LM-based); results are shown in Fig. 6. Considering the LM-based design, the number of available subcarriers increases in this way from $52 - 6 = 46$ to $58 - 8 = 50$.

The third example deals with the IEEE 802.11af standard for spectrum sharing among unlicensed white space devices and licensed services in TV bands [36], [37], and in particular the so-called "TV High Throughput (TVHT) W+W" configuration, in which two non-contiguous basic channel units (BCUs) are bonded. We consider a 6-MHz BCU setting, with subcarrier spacing $\Delta f = \frac{6}{144}$ kHz and $1/8$ CP overhead. The 802.11af standard specifies 104 used subcarriers per BCU, resulting in $\frac{104}{144} \approx 72.2\%$ spectral utilization. However, with the aid of the proposed precoders, it is possible to pack more subcarriers while still meeting the SEM constraints. Fig. 7 shows the results obtained with 144 used subcarriers per BCU.

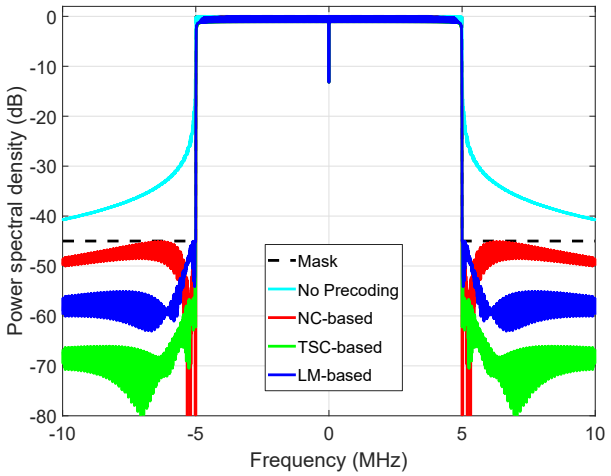

Fig. 8: PSD of various designs for 5G NR ACLR requirement of 45 dB.

The redundancy of the orthogonal precoders is in this case $R = 13$ (TSC-based), $R = 10$ (NC-based) and $R = 8$ (LM-based), leading to spectral utilization of 95.49%, 96.53% and 97.22% respectively.

Finally, we consider an LTE/5G NR scenario with a 10-MHz channel, subcarrier spacing $\Delta f = 15$ kHz and $9/128$ CP overhead. LTE allows a maximum of 50 resource blocks (groups of 12 subcarriers) in this configuration, leading to a spectral utilization of 90%. 5G NR is more aggressive in this aspect, allowing up to 52 resource blocks (93.6% utilization) [6]. Targeting maximum bandwidth utilization, we applied the orthogonal designs in this scenario with 666 used subcarriers, and assuming the Base Station Adjacent Channel Leakage Ratio (ACLR) limit of 45 dB as specified in [6]. Results are shown in Fig. 8: redundancies are $R = 12$ (TSC-based), $R = 10$ (NC-based) and $R = 9$ (LM-based), respectively yielding spectral utilizations of 98.1%, 98.4% and 98.55%.

VII. Conclusions

The mask-compliant orthogonal precoders presented conform to SEM requirements seeking minimal redundancy, which is beneficial in terms of throughput and on-line implementation complexity. Whereas the LM-based design is seen to allow for smaller redundancy than the NC- and TSC-based designs, the off-line computational cost can be significant, particularly for systems with a large number of subcarriers. Although this is not a problem for systems with static SEM requirements, it may favor the NC- and TSC-based approaches in CR scenarios with dynamic spectral masks in which the precoding matrix has to be recomputed over time.

Appendix
Proof of Theorem 1

To prove the first part, consider the derivative of the PSD $P(f_m) = \sum_{\ell=R}^{K-1} \mathbf{g}_\ell^H \mathbf{A}(f_m)\mathbf{g}_\ell$ with respect to μ_m, given by

$$\begin{aligned}\frac{\partial P(f_m)}{\partial \mu_m} &= \sum_{\ell=R}^{K-1} \left[\left(\mathbf{g}'_\ell\right)^H \mathbf{A}(f_m)\mathbf{g}_\ell + \mathbf{g}_\ell^H \mathbf{A}(f_m)\mathbf{g}'_\ell \right] \\ &= 2\,\mathrm{Re}\left\{ \sum_{\ell=R}^{K-1} \mathbf{g}_\ell^H \mathbf{A}(f_m)\mathbf{g}'_\ell \right\}, \end{aligned} \tag{27}$$

where $(\cdot)'$ indicates component-wise derivative with respect to μ_m, and with the last step following from $\mathbf{A}(f_m)$ being Hermitian. To obtain (27), we must find the derivative of the eigenvectors $\mathbf{g}_\ell$. Taking derivatives in $\mathbb{A}\mathbf{G} = \mathbf{G}\mathbf{\Lambda}$, one has

$$\mathbb{A}'\mathbf{G} - \mathbf{G}\mathbf{\Lambda}' = \mathbf{G}'\mathbf{\Lambda} - \mathbb{A}\mathbf{G}', \tag{28}$$

Multiplying both sides of (28) from the left by $\mathbf{G}^H$ and using the fact that $\mathbb{A} = \mathbf{G}\mathbf{\Lambda}\mathbf{G}^H$,

$$\mathbf{G}^H\mathbb{A}'\mathbf{G} - \mathbf{\Lambda}' = \mathbf{C}\mathbf{\Lambda} - \mathbf{\Lambda}\mathbf{C}, \tag{29}$$

where $\mathbf{C} \triangleq \mathbf{G}^H\mathbf{G}'$. If all eigenvalues are distinct, from (29) one has that the (k,ℓ) entry of $\mathbf{C}$ is given by

$$c_{k\ell} = \frac{\mathbf{g}_k^H \mathbb{A}' \mathbf{g}_\ell}{\lambda_\ell - \lambda_k}, \qquad \forall\, k,\ell = 0,1,\ldots,K-1, \quad k \neq \ell. \tag{30}$$

Since $\mathbf{G}' = \mathbf{G}\mathbf{C}$, the derivative of the ℓ-th eigenvector is

$$\begin{aligned}\mathbf{g}'_\ell &= \sum_{k=0}^{K-1} c_{k\ell}\mathbf{g}_k \\ &= c_{\ell\ell}\mathbf{g}_\ell + \sum_{k\neq\ell} \mathbf{g}_k \frac{\mathbf{g}_k^H \mathbb{A}' \mathbf{g}_\ell}{\lambda_\ell - \lambda_k} \\ &= c_{\ell\ell}\mathbf{g}_\ell + \left(\sum_{k\neq\ell} \frac{\mathbf{g}_k \mathbf{g}_k^H}{\lambda_\ell - \lambda_k} \right) \mathbb{A}'\mathbf{g}_\ell. \end{aligned} \tag{31}$$

Notice now that

$$\begin{aligned}\sum_{\substack{k=0 \\ k\neq\ell}}^{K-1} \frac{\mathbf{g}_k \mathbf{g}_k^H}{\lambda_\ell - \lambda_k} &= \mathbf{G}(\lambda_\ell \mathbf{I}_K - \mathbf{\Lambda})^\dagger \mathbf{G}^H \\ &= (\lambda_\ell \mathbf{I}_K - \mathbb{A})^\dagger, \end{aligned} \tag{32}$$

where we have used the pseudoinverse

$$(\lambda_\ell \mathbf{I}_K - \mathbf{\Lambda})^\dagger = \mathrm{diag}\left\{ \frac{1}{\lambda_\ell - \lambda_0} \quad \frac{1}{\lambda_\ell - \lambda_1} \quad \cdots \quad 0 \quad \cdots \quad \frac{1}{\lambda_\ell - \lambda_{K-1}} \right\} \tag{33}$$

where the zero is in the ℓ-th position. Therefore, (31) becomes

$$\mathbf{g}'_\ell = c_{\ell\ell}\mathbf{g}_\ell + (\lambda_\ell \mathbf{I}_K - \mathbb{A})^\dagger \mathbb{A}' \mathbf{g}_\ell. \tag{34}$$

Substituting (34) in (27),

$$\begin{aligned}\frac{\partial P(f_m)}{\partial \mu_m} &= 2\,\mathrm{Re}\left\{ \sum_{\ell=R}^{K-1} c_{\ell\ell} \mathbf{g}_\ell^H \mathbf{A}(f_m)\mathbf{g}_\ell \right\} \\ &\quad + 2\,\mathrm{Re}\left\{ \sum_{\ell=R}^{K-1} \mathbf{g}_\ell^H \mathbf{A}(f_m)(\lambda_\ell \mathbf{I}_K - \mathbb{A})^\dagger \mathbb{A}' \mathbf{g}_\ell \right\}. \end{aligned} \tag{35}$$

Taking derivatives in $\mathbf{G}^H\mathbf{G} = \mathbf{I}_K$, it is seen that $\mathbf{C} + \mathbf{C}^H = \mathbf{0}$, so that the diagonal entries $c_{\ell\ell}$ are purely imaginary. This means that the first term in (35) is zero, since $\mathbf{g}_\ell^H \mathbf{A}(f_m)\mathbf{g}_\ell$ is real-valued. On the other hand, since $\mathbb{A}' = \mathbf{A}(f_m)$, and both $\mathbb{A}$, $\mathbf{A}(f_m)$ are Hermitian, (35) boils down to

$$\begin{aligned}\frac{\partial P(f_m)}{\partial \mu_m} &= 2\sum_{\ell=R}^{K-1} \mathbf{g}_\ell^H \mathbf{A}(f_m)(\lambda_\ell \mathbf{I}_K - \mathbb{A})^\dagger \mathbf{A}(f_m)\mathbf{g}_\ell \\ &= 2\sum_{\ell=R}^{K-1}\sum_{\substack{k=0\\k\neq\ell}}^{K-1} \frac{|\mathbf{g}_\ell^H \mathbf{A}(f_m)\mathbf{g}_k|^2}{\lambda_\ell - \lambda_k} \qquad (36)\\ &= 2\sum_{\ell=R}^{K-1}\sum_{k=0}^{R-1} \frac{|\mathbf{g}_\ell^H \mathbf{A}(f_m)\mathbf{g}_k|^2}{\lambda_\ell - \lambda_k} \\ &\quad + 2\sum_{\ell=R}^{K-1}\sum_{\substack{k=R\\k\neq\ell}}^{K-1} \frac{|\mathbf{g}_\ell^H \mathbf{A}(f_m)\mathbf{g}_k|^2}{\lambda_\ell - \lambda_k}, \qquad (37)\end{aligned}$$

where in (36) we have used (32). The first term in (37) is nonpositive, since $\lambda_\ell - \lambda_k < 0$ for $\ell > k$; whereas the second term in (37) is zero because the (k,ℓ) term in the summation equals the negative of the (ℓ,k) term. The derivative sought can be written compactly by noting that $\mathbf{A}(f_m) = \mathbf{a}(f_m)\mathbf{a}^H(f_m)$, and defining

$$\gamma_\ell^m \triangleq \mathbf{g}_\ell^H \mathbf{a}(f_m), \qquad (38)$$

so that

$$\frac{\partial P(f_m)}{\partial \mu_m} = 2\sum_{k=0}^{R-1}\sum_{\ell=R}^{K-1} \frac{|\gamma_\ell^m|^2 |\gamma_k^m|^2}{\lambda_\ell - \lambda_k} \leq 0, \qquad (39)$$

which proves the first part of the theorem.

To prove the second part, let us define

$$\mathbb{A}_{\not m} \triangleq \sum_{\substack{n=0\\n\neq m}}^{D-1} \mu_n \mathbf{A}(f_n) = \mathbb{A} - \mu_m \mathbf{A}(f_m). \qquad (40)$$

Then we can write $Q(f_{\not m}) = \sum_{\ell=R}^{K-1} \mathbf{g}_\ell^H \mathbb{A}_{\not m} \mathbf{g}_\ell$, so that

$$\begin{aligned}\frac{\partial Q(f_{\not m})}{\partial \mu_m} &= 2\,\mathrm{Re}\left\{\sum_{\ell=R}^{K-1} \mathbf{g}_\ell^H \mathbb{A}_{\not m} \mathbf{g}_\ell'\right\} \\ &= 2\,\mathrm{Re}\left\{\sum_{\ell=R}^{K-1} \mathbf{g}_\ell^H \mathbb{A} \mathbf{g}_\ell'\right\} \\ &\quad - \mu_m \cdot 2\,\mathrm{Re}\left\{\sum_{\ell=R}^{K-1} \mathbf{g}_\ell^H \mathbf{A}(f_m)\mathbf{g}_\ell'\right\} \\ &= 2\sum_{\ell=R}^{K-1} \mathrm{Re}\left\{\mathbf{g}_\ell^H \mathbb{A}(c_{\ell\ell}\mathbf{g}_\ell + (\lambda_\ell \mathbf{I}_K - \mathbb{A})^\dagger \mathbf{A}(f_m)\mathbf{g}_\ell)\right\} \\ &\quad - \mu_m \frac{\partial P(f_m)}{\partial \mu_m} \qquad (41)\\ &= 2\sum_{\ell=R}^{K-1} \underbrace{\mathrm{Re}\left\{c_{\ell\ell}\mathbf{g}_\ell^H \mathbb{A}\mathbf{g}_\ell\right\}}_{=0} \\ &\quad + 2\sum_{\ell=R}^{K-1} \mathrm{Re}\left\{\mathbf{g}_\ell^H \mathbb{A}(\lambda_\ell \mathbf{I}_K - \mathbb{A})^\dagger \mathbf{A}(f_m)\mathbf{g}_\ell\right\} \\ &\quad - \mu_m \frac{\partial P(f_m)}{\partial \mu_m}. \qquad (42)\end{aligned}$$

Note now that, using $\mathbb{A} = \mathbf{G}\mathbf{\Lambda}\mathbf{G}^H$ and (32), one has

$$\begin{aligned}\mathbf{g}_\ell^H \mathbb{A}(\lambda_\ell \mathbf{I}_K - \mathbb{A})^\dagger &= \mathbf{g}_\ell^H \mathbf{G}\mathbf{\Lambda}(\lambda_\ell \mathbf{I}_K - \mathbf{\Lambda})^\dagger \mathbf{G}^H \\ &= \lambda_\ell \mathbf{e}_\ell^H (\lambda_\ell \mathbf{I}_K - \mathbf{\Lambda})^\dagger \mathbf{G}^H \\ &= 0, \qquad (43)\end{aligned}$$

in view of (33). Hence, the second term in (42) is zero, so that

$$\frac{\partial Q(f_{\not m})}{\partial \mu_m} = -\mu_m \frac{\partial P(f_m)}{\partial \mu_m} \geq 0, \qquad (44)$$

which concludes the proof of the second part.

Finally, using the definition $Q(f_{\not m}) = \sum_{n\neq m} \mu_n P(f_n)$ and taking derivatives, one has

$$\begin{aligned}\frac{\partial Q(f_{\not m})}{\partial \mu_m} &= \sum_{\substack{n=0\\n\neq m}}^{D-1} \mu_n \frac{\partial P(f_n)}{\partial \mu_m} \\ &= \sum_{n=0}^{D-1} \mu_n \frac{\partial P(f_n)}{\partial \mu_m} - \mu_m \frac{\partial P(f_m)}{\partial \mu_m}, \qquad (45)\end{aligned}$$

which, together with (44), proves the third part.

REFERENCES

[1] Report ITU-R SM.2421-0, "Unwanted emissions of digital radio systems," Jun. 2018.

[2] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.,* vol. 23, no. 2, pp. 201-220, Feb. 2005.

[3] M. Ibnkahla, *Cooperative Cognitive Radio Networks: The Complete Spectrum Cycle.* CRC Press: Boca Raton, FL, 2014.

[4] T. A. Weiss, F. K. Jondral, "Spectrum pooling: an innovative strategy for the enhancement of spectrum efficiency," *IEEE Commun. Mag., Radio Commun. Suppl.*, pp. 8-14, Mar. 2004.

[5] *Standard for Wireless Regional Area Networks (WRAN) – Specific Requirements – Part 22: Cognitive Wireless RAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Policies and procedures for operation in the TV Bands,* IEEE Standard 802.22, 2011.

[6] 3GPP TS 38.104 version 15.2.0 Release 15, 5G; NR; Base Station (BS) radio transmission and reception, Jul. 2018.

[7] B. Farhang-Boroujeny, "OFDM versus filter bank multicarrier," *IEEE Signal Process. Mag.,* vol. 28 no. 3, pp. 92-112, May 2011.

[8] A. A. Zaidi, R. Baldemair, H. Tullberg,H. Bjorkegren, L. Sundstrom, J. Medbo, C. Kilinc and I. Da Silva, "Waveform and numerology to support 5G services and requirements," *IEEE Commun. Mag.,* vol. 54 no. 11, pp. 90-98, Nov. 2016.

[9] M. Faulkner, "The effect of filtering on the performance of OFDM systems," *IEEE Trans. Veh. Technol.,* vol. 49 no. 5, pp. 1877-1884, Sep. 2000.

[10] T. Weiss, J. Hillenbrand, A. Krohn, F. K. Jondral, "Mutual interference in OFDM-based spectrum pooling systems," in *Proc. IEEE Veh. Technol. Conf.,* vol. 4, pp. 1873-1877, May 2004.

[11] R. Zayani, Y. Medjahdi, H. Shaiek, D. Roviras, "WOLA-OFDM: A potential candidate for asynchronous 5G," in *IEEE Global Commun. Workshops,* Dec. 2016.

[12] K. Hussain, R. López-Valcarce, "Optimal window design for W-OFDM," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.,* May 2020.

[13] H. A. Mahmoud, H. Arslan, "Sidelobe suppression in OFDM-based spectrum sharing system using adaptive symbol transition," *IEEE Commun. Lett.*, vol. 12, no. 2, pp. 113-135, Feb. 2008.

[14] H. Yamaguchi, "Active interference cancellation technique for MB-OFDM cognitive radio," in *Proc. Eur. Microw. Conf.,* vol. 2, pp. 1105-1108, Oct. 2004.

[15] S. Brandes, I. Cosovic, M. Schnell, "Reduction of out-of-band radiation in OFDM systems by insertion of cancellation carriers," *IEEE Commun. Lett.*, vol. 10, no. 6, pp. 420-422, June 2006.

[16] D. Qu, Z. Wang, T. Jiang, "Extended active interference cancellation for sidelobe suppression in cognitive radio OFDM systems with cyclic prefix," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1689-1695, May 2010.

[17] J. F. Schmidt, S. C. Sanz, R. López-Valcarce, "Choose your subcarriers wisely: active interference cancellation for cognitive OFDM," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 3, no. 4, pp. 615-625, Dec. 2013.

[18] C.-D. Chung, "Spectral precoding for rectangularly pulsed OFDM," *IEEE Trans. Commun.,* vol. 56, no. 9, pp. 1498-1510, Sep. 2008.

[19] R. Xu and M. Chen, "A precoding scheme for DFT-based OFDM to suppress sidelobes," *IEEE Commun. Lett.*, vol. 13. no. 10, pp. 776-778, Oct. 2009.

[20] J. van de Beek, F. Berggren, "*N*-continuous OFDM," *IEEE Commun. Lett.,* vol. 13, no. 1, pp. 1–3, Jan. 2009.

[21] J. van de Beek, "Sculpting the multicarrier spectrum: a novel projection precoder," *IEEE Commun. Lett.,* vol. 13, no. 12, pp. 881–883, Dec. 2009.

[22] M. Ma, X. Huang, B. Jiao, Y. J. Guo, "Optimal orthogonal precoding for power leakage suppression in DFT-based systems," *IEEE Trans. Commun.,* vol. 59, no. 3, pp. 844–853, Mar. 2011.

[23] J. Zhang, X. Huang, A. Cantoni, Y. J. Guo, "Sidelobe suppression with orthogonal projection for multicarrier systems," *IEEE Trans. Commun.,* vol. 60, no. 2, pp. 589–599, Feb. 2012.

[24] A. Tom, A. Sahin, H. Arslan, "Mask compliant precoder for OFDM spectrum shaping," *IEEE Commun. Lett.*, vol. 17, pp. 447–450, March 2013.

[25] S. Kant, G. Fodor, M. Bengtsson, B. Goransson, C. Fischione, "Low-complexity OFDM spectral precoding," *IEEE Int. Workshop Signal Process. Adv. Wireless Commun.,* 2019.

[26] R. Kumar, A. Tyagi, "Computationally-efficient mask-compliant spectral precoder for OFDM cognitive radio," *IEEE Trans. Cognitive Commun. Networking*, vol. 2, no. 1, pp. 15-23, March 2016.

[27] J. van de Beek, "Orthogonal multiplexing in a subspace of frequency well-localized signals," *IEEE Commun. Lett.*, vol. 14. no. 10, pp. 882–884, Oct. 2010.

[28] I. V. L. Clarkson, "Orthogonal precoding for sidelobe suppression in DFT-based systems using block reflectors," *IEEE Int. Conf. Acoust., Speech, Signal Process.,* 2017.

[29] X. Zhou, G. Y. Li, G. Sun, "Multiuser spectral precoding for OFDM-based cognitive radio systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 345–352, Mar. 2013.

[30] P. Bonami, M. Kilinç, J. Linderoth, "Algorithms and software for convex mixed integer nonlinear programs," in J. Lee, S. Leyffer (Eds.), *Mixed Integer Nonlinear Programming,* IMA Volumes in Mathematics and its Applications, vol. 154, pp. 1–40, Springer, Berlin, 2011.

[31] S. Burer, A. N. Letchford, "Non-convex mixed-integer nonlinear programming: A survey," *Surveys in Operations Research and Management Science,* vol. 17, no. 2, pp. 97–106, Jul. 2012.

[32] E. Kokiopoulou, J. Chen, Y. Saad, "Trace optimization and eigenproblems in dimension reduction methods," *Numerical Linear Algebra with Applications,* vol. 18, no. 3, pp. 565–602, May 2011.

[33] J. von Neumann, "Some matrix inequalities and metrization of matrix-space," *Tomsk. Univ. Rev.,* vol. 1, pp. 286–300, 1937.

[34] A. Mody, A. Saha, I. Reede, G. Miele, G. Cerro, "IEEE 802.22/802.22.3 Cognitive Radio Standards: Theory to Implementation." In: Zhang W. (eds), *Handbook of Cognitive Radio.* Springer, Singapore, 2017.

[35] IEEE 802 Standard; Part 11; Amendment 6: Wireless access in vehicular environments, IEEE, Jul. 2010.

[36] IEEE 802 Standard; Part 11; Amendment 5: Television White Spaces (TVWS) operation, IEEE, Feb. 2014.

[37] A. B. Flores, R. E. Guerra, E. W. Knightly, P. Ecclesine, S. Pandey, "IEEE 802.11af: a standard for TV white space spectrum sharing," *IEEE Commun. Mag.,* vol. 51 no. 10, pp. 92-100, Oct. 2013.